

PCT09

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/743,823

DATE: 11/14/2001

TIME: 14:15:45

Input Set : A:\6184844j.app

Output Set: N:\CRF3\11142001\I743823.raw

3 <110> APPLICANT: FOGHER, CORRADO
 5 <120> TITLE OF INVENTION: A SYNTHETIC POLYNUCLEOTIDE CODING FOR HUMAN
 6 LACTOFERRIN, VECTORS, CELLS AND TRANSGENIC PLANTS
 7 CONTAINING IT
 9 <130> FILE REFERENCE: 618484-4/JP/B-4075PCT
 11 <140> CURRENT APPLICATION NUMBER: 09/743,823
 12 <141> CURRENT FILING DATE: 2001-01-16
 14 <150> PRIOR APPLICATION NUMBER: IT RM98A000478
 15 <151> PRIOR FILING DATE: 1998-07-17
 17 <160> NUMBER OF SEQ ID NOS: 25
 19 <170> SOFTWARE: PatentIn Ver. 2.1
 21 <210> SEQ ID NO: 1
 22 <211> LENGTH: 2079
 23 <212> TYPE: DNA
 24 <213> ORGANISM: Artificial Sequence ✓
 26 <220> FEATURE:
 27 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic DNA ✓
 28 encoding human lactoferrin
 30 <220> FEATURE:
 31 <221> NAME/KEY: CDS
 32 <222> LOCATION: (1)..(2076)
 34 <400> SEQUENCE: 1

35	ggc	cgt	agg	aga	agg	agt	gtt	caa	tgg	tgc	gca	gta	tca	caa	cca	gag	48
36	Gly	Arg	Arg	Arg	Arg	Ser	Val	Gln	Trp	Cys	Ala	Val	Ser	Gln	Pro	Glu	
37	1				5					10					15		
39	gcc	aca	aaa	tgc	ttc	caa	tgg	caa	agg	aat	atg	aga	aaa	gtt	cgt	gga	96
40	Ala	Thr	Lys	Cys	Phe	Gln	Trp	Gln	Arg	Asn	Met	Arg	Lys	Val	Arg	Gly	
41				20					25					30			
43	cct	cct	gta	tct	tgc	ata	aag	aga	gat	tca	ccc	atc	cag	tgt	atc	cag	144
44	Pro	Pro	Val	Ser	Cys	Ile	Lys	Arg	Asp	Ser	Pro	Ile	Gln	Cys	Ile	Gln	
45			35					40					45				
47	gca	att	gcg	gaa	aac	aga	gct	gat	gct	gtg	act	ctt	gat	ggt	ggt	ttc	192
48	Ala	Ile	Ala	Glu	Asn	Arg	Ala	Asp	Ala	Val	Thr	Leu	Asp	Gly	Gly	Phe	
49		50					55					60					
51	ata	tac	gag	gca	gga	ctt	gcc	cca	tac	aaa	ctg	cga	cct	gta	gcg	gcg	240
52	Ile	Tyr	Glu	Ala	Gly	Leu	Ala	Pro	Tyr	Lys	Leu	Arg	Pro	Val	Ala	Ala	
53	65					70					75					80	
55	gaa	gtc	tac	ggg	acc	gaa	aga	caa	cca	cga	act	cac	tat	tat	gct	gtg	288
56	Glu	Val	Tyr	Gly	Thr	Glu	Arg	Gln	Pro	Arg	Thr	His	Tyr	Tyr	Ala	Val	
57				85						90				95			
59	gct	gtt	gtg	aag	aag	ggc	gga	tct	ttt	cag	ctg	aac	gaa	ctt	caa	ggt	336
60	Ala	Val	Val	Lys	Lys	Gly	Gly	Ser	Phe	Gln	Leu	Asn	Glu	Leu	Gln	Gly	
61				100					105					110			
63	ctg	aag	tca	tgc	cac	aca	gga	ctt	cgc	agg	acc	gct	gga	tgg	aat	gtc	384
64	Leu	Lys	Ser	Cys	His	Thr	Gly	Leu	Arg	Arg	Thr	Ala	Gly	Trp	Asn	Val	
65			115				120						125				
67	cct	ata	ggg	aca	ctt	cgt	cca	ttc	ttg	aat	tgg	acg	ggt	cca	cct	gag	432

ENTERED

P.5

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/743,823

DATE: 11/14/2001

TIME: 14:15:45

Input Set : A:\6184844j.app

Output Set: N:\CRF3\11142001\I743823.raw

```

68 Pro Ile Gly Thr Leu Arg Pro Phe Leu Asn Trp Thr Gly Pro Pro Glu
69      130      135      140
71 ccc att gag gca gct gtg gca aga ttc ttc tca gcc tct tgt gtt cca 480
72 Pro Ile Glu Ala Ala Val Ala Arg Phe Phe Ser Ala Ser Cys Val Pro
73 145      150      155      160
75 ggt gca gat aaa gga caa ttc ccc aac ctt tgt cgc ctg tgt gcg ggg 528
76 Gly Ala Asp Lys Gly Gln Phe Pro Asn Leu Cys Arg Leu Cys Ala Gly
77      165      170      175
79 aca ggg gaa aac aaa tgt gca ttc tca tcc cag gaa ccg tac ttc agc 576
80 Thr Gly Glu Asn Lys Cys Ala Phe Ser Ser Gln Glu Pro Tyr Phe Ser
81      180      185      190
83 tac tct ggt gcc ttt aag tgt ctt aga gac ggt gct gga gat gtt gct 624
84 Tyr Ser Gly Ala Phe Lys Cys Leu Arg Asp Gly Ala Gly Asp Val Ala
85      195      200      205
87 ttt att aga gag agc aca gtg ttt gag gat ctt tca gac gag gct gaa 672
88 Phe Ile Arg Glu Ser Thr Val Phe Glu Asp Leu Ser Asp Glu Ala Glu
89      210      215      220
91 agg gac gag tat gag tta ctc tgc cca gac aac act cgt aag cca gtt 720
92 Arg Asp Glu Tyr Glu Leu Leu Cys Pro Asp Asn Thr Arg Lys Pro Val
93 225      230      235      240
95 gac aag ttc aaa gat tgc cat ctt gca cgg gtc cct tct cat gcc gtt 768
96 Asp Lys Phe Lys Asp Cys His Leu Ala Arg Val Pro Ser His Ala Val
97      245      250      255
99 gtg gca cga agt gtt aat gga aag gag gat gcc atc tgg aat ctt ctc 816
100 Val Ala Arg Ser Val Asn Gly Lys Glu Asp Ala Ile Trp Asn Leu Leu
101      260      265      270
103 cgc caa gca cag gaa aag ttt gga aag gac aag tca ccg aaa ttc cag 864
104 Arg Gln Ala Gln Glu Lys Phe Gly Lys Asp Lys Ser Pro Lys Phe Gln
105      275      280      285
107 ctc ttt ggt tcc cct agt ggg cag aaa gat ctt ctg ttc aag gac tct 912
108 Leu Phe Gly Ser Pro Ser Gly Gln Lys Asp Leu Leu Phe Lys Asp Ser
109      290      295      300
111 gcc att ggg ttt tcg aga gtg cca cct agg ata gat tct ggg ttg tac 960
112 Ala Ile Gly Phe Ser Arg Val Pro Pro Arg Ile Asp Ser Gly Leu Tyr
113 305      310      315      320
115 ctt ggc tcc gga tac ttt act gca att cag aac ttg agg aaa agt gag 1008
116 Leu Gly Ser Gly Tyr Phe Thr Ala Ile Gln Asn Leu Arg Lys Ser Glu
117      325      330      335
119 gag gaa gtt gct gcc cgg cgt gcg cgg gtc gtt tgg tgt gcg gtg gga 1056
120 Glu Glu Val Ala Ala Arg Arg Ala Arg Val Val Trp Cys Ala Val Gly
121      340      345      350
123 gag caa gag ttg cgc aag tgt aac cag tgg agt ggt ttg agc gaa gga 1104
124 Glu Gln Glu Leu Arg Lys Cys Asn Gln Trp Ser Gly Leu Ser Glu Gly
125      355      360      365
127 tct gtg acc tgc tca tcg gcc tcc act aca gaa gat tgc atc gcc ctg 1152
128 Ser Val Thr Cys Ser Ser Ala Ser Thr Thr Glu Asp Cys Ile Ala Leu
129      370      375      380
131 gtg ttg aaa gga gaa gct gat gcc atg agt ttg gat gga gga tat gtt 1200
132 Val Leu Lys Gly Glu Ala Asp Ala Met Ser Leu Asp Gly Gly Tyr Val

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/743,823

DATE: 11/14/2001

TIME: 14:15:45

Input Set : A:\6184844j.app

Output Set: N:\CRF3\11142001\I743823.raw

133	385				390				395				400				
135	tac	act	gca	ggt	aaa	tgt	ggt	ttg	gtg	cct	gtc	ctt	gca	gag	aac	tac	1248
136	Tyr	Thr	Ala	Gly	Lys	Cys	Gly	Leu	Val	Pro	Val	Leu	Ala	Glu	Asn	Tyr	
137					405				410						415		
139	aaa	tca	caa	caa	agc	agt	gac	cct	gat	cct	aac	tgt	gtg	gat	aga	cct	1296
140	Lys	Ser	Gln	Gln	Ser	Ser	Asp	Pro	Asp	Pro	Asn	Cys	Val	Asp	Arg	Pro	
141					420				425						430		
143	gtg	gaa	gga	tat	ctt	gct	gtg	gcg	gtg	gtt	agg	aga	tca	gac	act	agc	1344
144	Val	Glu	Gly	Tyr	Leu	Ala	Val	Ala	Val	Val	Arg	Arg	Ser	Asp	Thr	Ser	
145					435				440						445		
147	ctt	acc	tgg	aac	tct	gtg	aaa	ggc	aag	aag	tcc	tgc	cac	acc	gcc	gtg	1392
148	Leu	Thr	Trp	Asn	Ser	Val	Lys	Gly	Lys	Lys	Ser	Cys	His	Thr	Ala	Val	
149					450				455						460		
151	gac	agg	act	gca	ggt	tgg	aat	atc	ccc	atg	gga	ttg	ctc	ttc	aac	cag	1440
152	Asp	Arg	Thr	Ala	Gly	Trp	Asn	Ile	Pro	Met	Gly	Leu	Leu	Phe	Asn	Gln	
153	465					470					475					480	
155	acg	ggc	tcc	tgc	aaa	ttt	gat	gaa	tat	ttc	agt	caa	agc	tgt	gcc	cct	1488
156	Thr	Gly	Ser	Cys	Lys	Phe	Asp	Glu	Tyr	Phe	Ser	Gln	Ser	Cys	Ala	Pro	
157					485					490					495		
159	ggt	tct	gac	cca	aga	tct	aat	ctc	tgt	gct	ttg	tgt	att	gga	gat	gag	1536
160	Gly	Ser	Asp	Pro	Arg	Ser	Asn	Leu	Cys	Ala	Leu	Cys	Ile	Gly	Asp	Glu	
161					500					505					510		
163	caa	ggt	gag	aat	aag	tgc	gtt	ccc	aac	agc	aac	gag	aga	tac	tac	ggt	1584
164	Gln	Gly	Glu	Asn	Lys	Cys	Val	Pro	Asn	Ser	Asn	Glu	Arg	Tyr	Tyr	Gly	
165					515					520					525		
167	tac	act	ggg	gct	ttc	cgt	tgc	ttg	gct	gag	aat	gct	gga	gac	gtt	gca	1632
168	Tyr	Thr	Gly	Ala	Phe	Arg	Cys	Leu	Ala	Glu	Asn	Ala	Gly	Asp	Val	Ala	
169					530					535					540		
171	ttt	gtg	aaa	gat	gtc	act	gtc	ttg	cag	aac	act	gat	gga	aat	aac	aat	1680
172	Phe	Val	Lys	Asp	Val	Thr	Val	Leu	Gln	Asn	Thr	Asp	Gly	Asn	Asn	Asn	
173	545					550				555					560		
175	gag	gca	tgg	gct	aag	gat	ttg	aag	ctt	gca	gac	ttt	gcg	ttg	ctg	tgc	1728
176	Glu	Ala	Trp	Ala	Lys	Asp	Leu	Lys	Leu	Ala	Asp	Phe	Ala	Leu	Leu	Cys	
177					565					570					575		
179	ctc	gat	ggc	aaa	cgt	aag	cct	gtg	act	gaa	gct	aga	agc	tgc	cat	ctt	1776
180	Leu	Asp	Gly	Lys	Arg	Lys	Pro	Val	Thr	Glu	Ala	Arg	Ser	Cys	His	Leu	
181					580					585					590		
183	gcc	atg	gcc	ccg	aat	cat	gct	gtg	gtg	tct	cgt	atg	gat	aag	gtg	gaa	1824
184	Ala	Met	Ala	Pro	Asn	His	Ala	Val	Val	Ser	Arg	Met	Asp	Lys	Val	Glu	
185					595					600					605		
187	cgc	ttg	aaa	cag	gtg	ttg	ctc	cac	caa	cag	gct	aaa	ttt	ggt	aga	aat	1872
188	Arg	Leu	Lys	Gln	Val	Leu	Leu	His	Gln	Gln	Ala	Lys	Phe	Gly	Arg	Asn	
189					610					615					620		
191	gga	tct	gac	tgc	ccg	gac	aag	ttt	tgc	tta	ttc	cag	tct	gaa	acc	aaa	1920
192	Gly	Ser	Asp	Cys	Pro	Asp	Lys	Phe	Cys	Leu	Phe	Gln	Ser	Glu	Thr	Lys	
193	625					630				635					640		
195	aac	ctt	ttg	ttc	aat	gac	aac	act	gag	tgt	ctt	gcc	aga	ctc	cat	ggc	1968
196	Asn	Leu	Leu	Phe	Asn	Asp	Asn	Thr	Glu	Cys	Leu	Ala	Arg	Leu	His	Gly	
197					645					650					655		

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/743,823

DATE: 11/14/2001

TIME: 14:15:45

Input Set : A:\6184844j.app

Output Set: N:\CRF3\11142001\I743823.raw

```

199 aaa aca aca tat gaa aaa tat ttg gga cca cag tat gtc gca ggc att 2016
200 Lys Thr Thr Tyr Glu Lys Tyr Leu Gly Pro Gln Tyr Val Ala Gly Ile
201          660          665          670
203 act aat ctg aaa aag tgc tca acc tcc cca ctc cta gaa gcc tgt gaa 2064
204 Thr Asn Leu Lys Lys Cys Ser Thr Ser Pro Leu Leu Glu Ala Cys Glu
205          675          680          685
207 ttc cta agg aag taa 2079
208 Phe Leu Arg Lys
209 690
212 <210> SEQ ID NO: 2
213 <211> LENGTH: 30
214 <212> TYPE: DNA
215 <213> ORGANISM: Artificial Sequence
217 <220> FEATURE:
218 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic DNA ✓
220 <400> SEQUENCE: 2
221 ggatccatgg gccgtaggag aaggagtgtt 30
224 <210> SEQ ID NO: 3
225 <211> LENGTH: 32
226 <212> TYPE: DNA
227 <213> ORGANISM: Artificial Sequence
229 <220> FEATURE:
230 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic DNA ✓
232 <400> SEQUENCE: 3
233 gagctccttc gggttttactt cctgaggaat tc 32
236 <210> SEQ ID NO: 4
237 <211> LENGTH: 42
238 <212> TYPE: DNA
239 <213> ORGANISM: Artificial Sequence
241 <220> FEATURE:
242 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic DNA ✓
244 <400> SEQUENCE: 4
245 tctagataaa ataatctata cattaaaaaa tttgatttta aa 42
248 <210> SEQ ID NO: 5
249 <211> LENGTH: 36
250 <212> TYPE: DNA
251 <213> ORGANISM: Artificial Sequence
253 <220> FEATURE:
254 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic DNA ✓
256 <400> SEQUENCE: 5
257 ggatccgact gagtcggata agaagaaaag aaaaga 36
260 <210> SEQ ID NO: 6
261 <211> LENGTH: 36
262 <212> TYPE: DNA
263 <213> ORGANISM: Artificial Sequence
265 <220> FEATURE:
266 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic DNA ✓
268 <400> SEQUENCE: 6
269 tctagagttt tcaaatttga attttaatgt gtgttg 36

```


RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/743,823

DATE: 11/14/2001

TIME: 14:15:45

Input Set : A:\6184844j.app

Output Set: N:\CRF3\11142001\I743823.raw

```

272 <210> SEQ ID NO: 7
273 <211> LENGTH: 36
274 <212> TYPE: DNA
275 <213> ORGANISM: Artificial Sequence
277 <220> FEATURE:
278 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic DNA
280 <400> SEQUENCE: 7
281 ggatcccacc ttaaggaggt tgcaacgagc gtggca 36
284 <210> SEQ ID NO: 8
285 <211> LENGTH: 250
286 <212> TYPE: DNA
287 <213> ORGANISM: Artificial Sequence
289 <220> FEATURE:
290 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic DNA
292 <400> SEQUENCE: 8
293 ggccgtagga gaaggagtgt tcaatgggtgc gcagtatcac aaccagaggc cacaaaatgc 60
294 ttccaatggc aaaggaatat gagaaaagtt cgtggacctc ctgtatcttg cataaagaga 120
295 gattcaccca tccagtgtat ccaggcaatt gcggaaaaca gagctgatgc tgtgactctt 180
296 gatggtggtt tcatatacga ggcaggactt gcccataca aactgcgacc tgtagcggcg 240
297 gaagtctacg 250
300 <210> SEQ ID NO: 9
301 <211> LENGTH: 250
302 <212> TYPE: DNA
303 <213> ORGANISM: Artificial Sequence
305 <220> FEATURE:
306 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic DNA
308 <400> SEQUENCE: 9
309 gcacctggaa cacaagaggc tgagaagaat cttgccacag ctgcctcaat gggctcaggt 60
310 ggacctgccc aattcaagaa tggacgaagt gtccctatag ggacattcca tccagcggtc 120
311 ctgcgaagtc ctgtgtggca tgacttcaga ccttgaagtt cgttcagctg aaaagatccg 180
312 cccttcttca caacagccac agcataatag tgagttcgtg gttgtctttc ggtcccgtag 240
313 acttccgccc 250
316 <210> SEQ ID NO: 10
317 <211> LENGTH: 250
318 <212> TYPE: DNA
319 <213> ORGANISM: Artificial Sequence
321 <220> FEATURE:
322 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic DNA
324 <400> SEQUENCE: 10
325 aactggctta cgagtgttgt ctgggcagag taactcatac tcgtcccttt cagcctcgtc 60
326 tgaaagatcc tcaaacactg tgctctctct aataaaagca acatctccag caccgtctct 120
327 aagacactta aaggcaccag agtagctgaa gtacggttcc tgggatgaga atgcacattt 180
328 gttttcccct gtccccgcac acaggcgaca aagggtgggg aattgtcctt tatctgcacc 240
329 tggaacacaa 250
332 <210> SEQ ID NO: 11
333 <211> LENGTH: 255
334 <212> TYPE: DNA
335 <213> ORGANISM: Artificial Sequence
337 <220> FEATURE:

```

Use of n and / or Xaa has been detected in the Sequence Listing. Review the Sequence Listing to ensure a corresponding explanation is present in the <220> to <223> fields of each sequence using n or Xaa.

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/743,823

DATE: 11/14/2001

TIME: 14:15:46

Input Set : A:\6184844j.app

Output Set: N:\CRF3\11142001\I743823.raw

L:758 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:25

L:762 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:25

L:763 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:25

L:764 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:25